Offending with Impunity: Racial Vilification and Freedom of Speech

WOJCIECH SADURSKI*

I. Introduction

Poles, as you all probably know, are stupid, lazy and utterly dishonest. They are all constant trouble-makers: a brief look at the history of this peculiar nation should convince everyone that all they have been doing throughout the ages was to conspire against the security and well-being of their neighbours. But this is their, and their neighbours', problem. However, once allowed to migrate to other countries such as this great country of ours, they immediately become a force which contributes to moral decay, criminality and subversion. You must have heard the one about a Polish omelette (how to make one? — first, you steal three eggs...) or about how many Poles it takes to replace a light bulb — well, this reflects something, doesn't it. Oh, and they never use deodorants.

The passage above, if taken seriously, and if included in a context which suggests that it is *meant* seriously, is likely to incite hatred towards, serious contempt for, or severe ridicule of a Polish ethnic minority.[1] As such, its publication would now be punishable in New South Wales, just as it would in much of the modern world. Indeed, the supreme international body resolved that the UN member states should "declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred".[2]

This in itself suggests that the matter is serious and deserves deeper consideration. First, it is serious politically in this country: while New South Wales and Western Australia are presently the only Australian states that prohibit public acts of racial vilification and incitement to racism,[3] there are moves in a similar direction in the other Australian states[4] as well as

---

* Reader in the Department of Jurisprudence, the University of Sydney. I am very grateful to Craig Carracher, Francine Feld, Kenneth Gergen, Michael Kirby, Martin Krygier, Chandran Kukathas, Wotek Lamentowicz, Patricia Loughlan, Sarah McNaughton, Andy Michels, Philip Pettit and David Tucker for their comments and suggestions.

1 My own use of the passage in this article exempts me, I hope, from these charges: morally — because I am a Pole myself, hence beyond the suspicion of harbouring such great self-hate; legally — because it is done here "reasonably and in good faith", "for academic purposes" which include "discussion or debate" about the law against racial vilification. The quoted words are taken from article (2)(c), *Anti-Discrimination (Racial Vilification) Amendment Act* NSW 1989.

2 *International Convention on the Elimination of All Forms of Racial Discrimination* (1965), Art 4 (a).

3 See, respectively, *Anti-Discrimination (Racial Vilification) Amendment Act* NSW 1989 s20D and *Criminal Code* 1913 (WA) ss76-80.

4 For example, in Victoria a special committee to advise the Attorney General on racial vilification released in June 1990 an Issues Paper (*Racial Vilification in Victoria*) for

proposals to entrench this prohibition at a federal level.[5] It is also serious internationally: the legal systems of the United Kingdom, Canada, New Zealand and most European states, as well as binding international human rights instruments, prohibit the dissemination of ideas hostile to racial (and other) groups.[6] The United States is an exception: so far, attempts there to entrench a similar prohibition have consistently been held unconstitutional.[7] Even there, however, demands are growing to abandon this non-interventionist approach.

Second, it is serious as a matter of legal doctrine: the offence of racial vilification intersects with a number of pre-existing torts or crimes (defamation, seditious libel, incitement to commit unlawful acts, intentional infliction of emotional distress, etc) without being identical with any of these. Moreover, the legal implications of the new offence are worth considering: they touch on matters central to the legal analysis of freedom of speech regulations. For example, what standard of scrutiny should apply in the judicial review of laws that restrict offensive speech? Should this standard of review depend on the absence of any discernible social value in the acts of speech?

Third, and most importantly for the purposes of this article, the matter has serious moral and philosophical implications. Most obviously problematic is its potential to conflict with the very foundations of a liberal and democratic state: freedom of expression. A widely accepted moral principle of such a state is that the very offensiveness of a given act of expression, and that it may cause some distress to other people, are insufficient bases for a legal prohibition. But surely it does not follow that, under liberalism, all acts of expression are absolutely pro- tected; there are generally unobjectionable laws which restrict freedom of speech for the purposes of protecting the administration of justice, national security, reputation and privacy. On which side of the divide should acts of group vilification fall? And why?

It seems clear that the proper context for a discussion about racial vilification laws is a philosophical one; we must reflect upon the justifications for, the nature of, and the limits to our commitment to freedom of expression. Only then can we give clear answers to questions about whether, and why, the dissemination of ideas hostile to groups should be prohibited. Hence, much of

---

comment; see "The Issue of Racial Vilification", *Law Institute Journal* (August 1990) at 709.

5 In 1984 the then Human Rights Commission recommended that amendments be made to the *Racial Discrimination Act* 1975 (Cth) to outlaw racial defamation and incitement to racial hatred, see Bailey, P, *Human Rights: Australia in an International Context* (1990) at 209-210. More recently, the Australian Law Reform Commission invited submissions on whether a federal offence prohibiting incitement to racial hatred should be created, see *Discussion Paper 48: Multiculturalism: Criminal Law* (Australian Law Reform Commission, Sydney 1991) 43-44 (par 4.29).

6 For a useful overview, see *Incitement to Racial Hatred: The International Experience*, Occasional Paper No 2, Australian Human Rights Commission (1982); Matsuda, M, "Public Response to Racist Speech: Considering the Victim's Story", *MichLRev* 87 (1989) at 2320, 2341-48. On treatment of group libel in international law, see Lerner, N, "Group Libel Revisited", *Israel Yearbook on Human Rights* 17 (1987) at 184, 195-6.

7 Most recently, see *Doe v University of Michigan*, 721 F Supp 852 (E D Mich 1989). The last case in which a similar act was upheld was *Beauharnais v Illinois*, 343 US 250 (1952). *Beauharnais* is generally held no longer to be good law, see *American Booksellers Association v Hudnut*, 771 F 2d 323, 331 n3 (7th Cir 1985).

this article will concern general issues in the philosophy of free speech. Without that, our inquiry into the specific question would be made in a moral vacuum.

I will proceed as follows. First, I will set the scene for my argument by setting the parameters of the concept of group vilification (part II). I will then reflect upon a legal framework for judging regulations of speech such as group vilification. In particular, I will discuss some proposed strategies in favour of lowering the usual stringent standard of testing such regulations which, more often than not, is fatal to restrictions of speech (part III). One conclusion of this part is that no such "lowering" of the standard of scrutiny can be made in isol- ation from the main purposes served by freedom of speech. Hence, in part IV, I examine those purposes and conclude that what provides the best justification for strong protection of freedom of speech is the function of political speech in maintaining the democratic process. Consequently, those kinds of speech that are relevant to debate about public issues may be restricted but only if they result in harm of particularly high gravity, and if the relationship between that speech and the resultant harm is sufficiently close. This leads directly to part V which con- tains an argument central to this article. It offers a list of the main harms prod- uced by group vilification, and (consistent with the conclusion of part IV) subjects these harms to a level of scrutiny appropriate to freedom of speech considerations. It is on the basis of these reflections (upon the nature of harm caused) that the conclusions about appropriateness (or otherwise) of legal restrictions upon group vilification are offered.

II. Defining Group Vilification

This part of my article will not so much offer a comparative analysis of various provisions that define racial vilification in international and national legal texts, but rather, will indicate the broad contours of the category for the purposes of further discussion. A useful way of doing this is to locate racial vilification on a continuum of types of offensive, dangerous or otherwise objectionable speech. At one extreme of this scale are the cases of speech that, even to the most com- mitted libertarian, would uncontroversially warrant restriction. At the other extreme are the cases of speech that (even to those who support legal prohibition of group vilification) obviously should not be legally restricted. By proceeding in a negative rather than a positive way (ie, by saying what group vilification *is not*) I hope to minimise the danger of smuggling in substantive conclusions through a definitional fiat.

We begin from the non-punishable end of the spectrum, that is, by listing those cases of potentially offensive speech where a consensus may be rather easily mobilised against any legal prohibition or punishment. First, it will probably be widely agreed that statements of verifiable truth should be legally protected even if a particular group may find them offensive. Second, the *context* of a statement may suggest that neither was it prompted by wrongful motives (such as the desire to vilify a group) nor that evil effects (such as offence caused) are likely to occur. One such example is the offensive passage that opens this article. More typically (but also more controversially) there will be cases of satire (but does that include racist jokes?), words uttered affectionately with no intention to hurt the hearer, a fair report of a public act,

or academic or scientific inquiry which may be distressing to some groups and yet is not motivated by hatred of those groups.

Third, no restriction is warranted when objections to an act of speech result from someone's excessive, unusual sensitivity. A measure of "unusual sensitivity" must be gauged by the normal sensitivity *of a target group* of the act of speech in question; otherwise it would be all too easy for a dominant social group to dismiss any complaint as a case of undue sensitivity.[8]

Fourth, acts of speech which are *private* rather than public should escape the threat of legal prohibition. It would perhaps be pedantic to ask "what is the public?" We may rather adopt a commonsensical approach that at least some types of speech are undoubtedly "private" (such as inaudibly murmuring words to oneself). And yet, any communication is by definition "public". So perhaps what matters is not so much the existence of an audience but rather the nature of the communicative behaviour: some types of behaviour are, by virtue of social conventions, private (for instance, a "private" dinner party conversation).[9] But how far should we rely on the dominant social conventions in defining "publicness"? One American university recently legislated in its anti-harassment regulations that, among other things, the following examples of race-based harassment are prohibited: "use of derogatory names, inappropriately directed laughter, . . . and conspicuous exclusion from conversations and/or classroom discussions".[10] Is "exclusion from conversations" the sort of communicative behaviour that is private or public?

These four categories seem to constitute (with all the reservations and doubts indicated) the most important acts of potentially offensive speech that should remain legally permissible, and they will not be taken into account here as putatively prohibited instances of group vilification. Now to the other end of the spectrum: instances of the kind of speech that is uncontroversially non-protected. This end of the spectrum contains two categories: incitement and defamation.

A number of provisions, both in international and national legal instruments, refer to "incitement to violence" alongside racial vilification.[11] And yet, incitement to violence, constituting as it does incitement to commit unlawful acts, is beyond the focus of this article since it does not entail any especially problematic issues from the point of view of freedom of speech. Of course it may be argued that acts of racial vilification constitute a sort of incitement to violence or other unlawful acts. But this is a matter of substantive argument which will be considered later, in part V(B). At this stage it suffices to say that if indeed a given act of racial vilification amounts

---

8 See Love, J C, "Discriminatory Speech and the Tort of Intentional Infliction of Emotional Distress", *Wash & Lee LRev* 47 (1990) at 123, 148-9.

9 The *Canadian Criminal Code* (1970), in its provisions dealing with "hate propaganda", implicitly exempts "private conversation" from an offence of "wilfully promoting hatred against any identifiable group", art 281.2(2).

10 Quoted in Post, R C,"Racist Speech, Democracy, and the First Amendment" *Wm & Mary LRev* 32 (1991) at 267, 269.

11 See for example, Article 20 of the *International Covenant on Civil and Political Rights* which declares that "any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law".

to the incitement of racially motivated violence, then there is no reason to resort to a separate crime of racial vilification.

The same applies to the commonly accepted indicia of defamation, as present in common law and various criminal codes. For all the differences between the laws of defamation in various legal systems, I will take it that prohibition of defamation is reasonably unproblematic when applied to "individual" targets of defamation. One general principle in the English and American laws of defamation is that it will be punishable only when targeting identifiable individuals. This is traditionally expressed in the principle that a defamatory statement is not actionable unless published "of and concerning" the plaintiff. Whether group vilification should also be punishable is a matter of separate substantive argument. Indeed, it may be claimed that group offences are more hurtful to individuals than "personalised" ones. I wish to defer consideration of this point until part V(C) of the article; for the moment, it suffices to say that nothing is gained by extending the notion of individual defamation to group vilification, and it is clear that there are many who would find the former, but not the latter, a proper object of legal concern.

The upshot of the preceding remarks is that, for the purposes of this discussion, a statement will not be "racially vilifying" if it is verifiably true; or is made in a context that suggests neither a malicious intention nor a likely damaging effect; or is made in private; or the resultant offence stems from unusual sensitivity of a member of a group. Furthermore, the concept of "racial vilification" as used here does not embrace statements that are defamatory in the sense of individual defamation, or constitute incitement to unlawful acts against a particular group.

III. What Level of Scrutiny?

There are words that hurt, and the words that express racial hate or contempt for an entire group of people hurt more than many other unpleasant words. It is one thing to say that offensiveness itself is an insufficient reason to punish the speaker, and that the regime of liberty demands that we put up with expressions of unwelcome ideas, unpopular thoughts and offensive views. It is an altogether different thing to claim that gratuitous harm, of whatever gravity, must always be tolerated merely because it results from words rather than from actions. Even the most committed liberal must accept that sometimes legislators "may create remedies for the damage done with words so long as these remedies display sufficient sensitivity to freedom of expression as well".[12]

The purpose of this and the following parts of this article is to place the argument about legal prohibition of racial vilification in a framework that facilitates legal reasoning about rights, and thus displays "sufficient sensitivity to freedom of expression". The idea is that one clear implication of talking about a "right" to free speech is that it insulates speech from the ordinary operation of the harm principle. By the "ordinary operation" of the harm principle I mean the use of a calculus of benefits (in terms of harm

12 Tribe, L H, *American Constitutional Law* (1988, 2nd edn), par 12-10 at 856: "The first amendment need not sanctify the deliberate infliction of pain simply because the vehicle used is verbal or symbolic rather than physical".

avoided) and costs (in terms of speech restrained). Under the calculus, the government would be obliged to prohibit a speech act whenever it would be reasonable to expect that the resultant harm outweighed the harm of restraining people from speaking. The regime of "rights", it is generally thought, insulates a given sphere of activity from the use of such a calculus: the notion of "rights" would be redundant if their exercise were conditional upon case-by-case confirmation that the benefits resulting from an individual action outweigh its costs. This does not mean that a right to act in a particular way makes *any* interference with such action inappropriate, but only that the threshold of justification must be substantially raised.

This much may be obvious: the question is how to translate it into language suited to legal argument. The framework offered here is in terms (familiar in American constitutionalism) of a level of scrutiny of a given regulation. The idea is to imagine yourself as a judge whose task it is to review a challenged regulation from the point of view of its conformity with a constitutional right (here: the right of free expression). A preliminary question that one must answer concerns the level of scrutiny: how great must a good be in order to warrant a restriction of the right, and how closely must the restriction be related to a likely achievement of this good? On one end of the spectrum of possible answers (lenient scrutiny) we may decide that any net surplus of benefit (however marginal) will justify restriction of a harmful action, and that even a slight likelihood of a positive relationship between the restriction and the benefit will do. On the opposite end of the spectrum (strict scrutiny), we may require a social good of very great importance to be cited in order to warrant a restriction, and a very strict relationship between the restriction and the achievement of this good. A decision concerning the precise point of the scale at which to locate our test may determine the results of the test: the closer we approach the strict-scrutiny end of the spectrum, the less likely the regulation is to survive. But we must have a rational method of determining the proper level of scrutiny.

This, and the following parts of this article will be devoted to an examination of such a method, in the context of prohibition of racial vilification. An introductory observation: while a framework of various levels of scrutiny has been elaborated in a system of judicial review of laws under a constitutional bill of rights, the underlying idea adopted here is that the framework is of universal application. As long as we presuppose that there is a *right* to free speech in a given society (whatever the sources of this right may be, constitutional or otherwise), we can analyse proposed restrictions on that right in the language of different levels of scrutiny.

A. Speech-plus and pure speech

One strategy of argument for a lowered level of scrutiny of speech acts is to detect a high "conduct ingredient" in a given act. Since it is only "speech" and not the "conduct" which, under the principle of freedom of expression, warrants a high degree of protection (over and above a minimum net surplus of benefit over harm), "speech-plus" will be entitled to a standard level of protected behaviour, ie, subject to normal operation of the harm principle.[13]

---

13 See, for example, *Teamsters Local 695 v Vogt*, 354 US 284, 289 (1957); *Cox v Louisiana* 379 US 559, at 562-4 (1965).

But the speech/conduct distinction is unhelpful for our subject matter for several reasons. First, any speech is accompanied by conduct of one sort or another, which may be burdensome to others: for example, talk may be noisy or the distribution of pamphlets may cause litter. This immediately suggests that the characterisation of an act as non-speech is a *result* of, not a prerequisite to, our willingness to regulate it.

Second, there are many forms of "conduct" the aim of which is primarily, or exclusively, communicative. Flag burning is a kind of "conduct" but its main aim is symbolic and communicative: that is, to convey a message that can be expressed in words, but which (according to a "speaker") will be more effective when the vehicle of symbolic action is used.[14] Consider, similarly, examples of displaying flags, wearing uniforms, armbands or signs on jackets, or saluting a flag. To protect speech, but not "conduct" that is equivalent to speech in the motives of an agent and/or in the meanings it actually conveys to the audience, seems inconsistent.[15]

Furthermore, not all instances of "pure speech" are necessarily considered "speech" for the purposes of the jurisprudence of free expression: it has often been pointed out that there are cases of "speech" in which the communicative component is so small (threatening phone calls, conspiracy to fix prices, "expressive entertainment", etc) that there is hardly any freedom-of-speech based rationale for special protection.[16] But this suggests that the very distinction between speech and "non-speech" is ultimately manipulable, and that if one wants to deny a high level of protection to a particular act, then one may simply define it as non-speech. Consequently, an appeal to this distinction appears unavailable to the proponents of group-vilification laws (if they wish to argue that the act is more than just speech) or to their opponents (if they insist that it is purely speech). Impunity of speech and of communicative conduct must stand or fall together.

B. Low value and high value speech

Another argument for lowering the level of scrutiny appeals to the social value of the speech in question. Racial-vilification speech has a lower social value than the kind of speech that a society has in mind when it entrenches rights to free expression. That not all speech is of equal importance from the point of view of First Amendment values has long been an established doctrine of the United States Supreme Court.[17] A list of "low value" categories of speech goes back to the landmark *Chaplinsky*[18] decision and,

---

14 See *Texas v Johnson*, 109 S Ct 2533, at 2539-40 (1989). The test, used in this flag-burning case, to decide whether a particular conduct expressing an idea may be labeled speech, is "whether '[a]n intent to convey a particularised message was present, and [whether] the likelihood was great that the message would be understood by those who viewed it'" (quoting the test articulated in *Spence v Washington*, 418 US 405, at 410-11 (1974)).

15 See Schauer, F, "Categories and the First Amendment: A Play in Three Acts", *Vanderbilt LRev* 34 (1981) 265, at 272-3; similarly, Posner, R, *The Problems of Jurisprudence* (1990) 467.

16 See Schauer, F, "Speech and 'Speech' — Obscenity and 'Obscenity': An Exercise in the Interpretation of Constitutional Language", *Georgetown LJ* 67 (1979) 899, at 905-910; Wright,R G, "A Rationale from J S Mill for the Free Speech Clause", *Supr Crt Rev* (1985) 149, at 163-9.

17 See *Dun & Bradstreet, Inc v Greenmoss Builders, Inc*, 472 US 749, at 758 (1985) ("not all speech is of equal First Amendment importance").

notwithstanding some recent dilutions,[19] includes obscenity, libel, express incitement, commercial speech, child pornography and "fighting words".

How much does this strategy help the proponent of lenient scrutiny of group vilification laws? In this context, to make a claim about the low value of a particular vilifying utterance may mean either of two things. First, it may mean that the harm produced by group vilification is significantly greater than in the case of many other categories of speech. But remember, the aim of a strategy for lowering the level of scrutiny of laws that restrict some types of speech is to justify a standard, non-suspicious (towards a legislator) test for the measurement of harm produced by a given speech. Hence, the reasons for lowering the level of the test themselves must be harm-independent. Otherwise, our appeal to a lower standard of scrutiny would collapse into the harm principle. But if *this* is all that is meant, the appeal to a lower standard of scrutiny is rendered redundant.

Alternatively, a "low-value" argument may mean that a given category of speech is so "remote" from the main rationale for freedom of expression that there are no moral or political reasons to accord it the usual freedom-of-expression protection.[20] After all, there are some *reasons* for having freedom of expression, and if these reasons do not apply to the category of speech in question, then there is no justification for treating it on a par with "core" categories of expression. Hence, in order to assess the weight of this argument, we must reflect upon the reasons for having freedom of expression in the first place, and only then test their application to the category of speech in question (here: racial vilification). It is only once we know what values are promoted by this freedom that we will be able to judge which categories of speech are "central", and which are "marginal", "penumbral" or "remote" (hence: of lower value). Rather than deciding this question here, we will postpone the discussion until the next part of this article.

C. *Viewpoint-based and viewpoint-neutral regulations*

The third strategy is different from the two previous ones in one major respect: it attaches the argument about a level of scrutiny not to the nature of speech, the restriction of which is to be scrutinised, but rather to the *motives* for a governmental restriction. The underlying idea is that the legitimacy of governmental prohibitions has to do with the nature of the motives for these prohibitions, rather than with the nature of the prohibited conduct itself. When applied to our subject-matter, the most influential version of the motive-related strategy refers to a distinction between viewpoint-based, as opposed to viewpoint-neutral, regulations of freedom of expression. The general principle might be formulated as follows: restrictions that express a viewpoint on a given issue that is preferred by the government (or conversely, that is disapproved of by the government) should be subjected to very strict scrutiny. By implication, restrictions that are viewpoint-neutral, and which do not discriminate among different viewpoints, may be subject to lenient

---

18 *Chaplinsky v New Hampshire*, 315 US 568 (1942).

19 See Tribe, note 12 above, par 12-10 at 850-56.

20 For an example of this approach, see Sunstein, C R, "Pornography and the First Amendment", *Duke LJ* (1986) 589.

scrutiny because there is little ground for suspicion that they result from improper governmental motives.[21]

In considering this strategy, one should distinguish between the postulate of viewpoint-neutrality and that of content-neutrality. It has sometimes been asserted that restrictions on freedom of expression should be not merely viewpoint-neutral but also content-neutral. This conclusion would follow from a literal reading of the principle announced by the United States Supreme Court: "the First Amendment means that government has no power to restrict expression because of its message, its ideas, its subject matter, or its content".[22] Literally read, this principle sounds wildly overstated, as it would lead to invalidation of even the most unobjectionable restrictions: is not a restriction of libel, child pornography, or misleading commercial advertising, based on the subject matter or the content of a message? The only plausible reading of the statement is that content-based restrictions should be subject to closer scrutiny than restrictions that are completely content-neutral (such as, for example, limits upon the size of roadside billboards regardless of their messages). Indeed, this appears to be a currently accepted doctrine of the United States Supreme Court and of much scholarly writing: content-based restrictions should trigger an intermediate level of scrutiny: between content-neutral ones (which warrant a very lenient test) and viewpoint-based ones (which trigger the strictest scrutiny).[23]

That the postulate of content-neutrality is different from, and much more radical than, the postulate of viewpoint-neutrality, is clear from its very formulation. What is even more important for present purposes is that the link between the demands for content-neutrality and the motive-based theory of legitimate state prohibitions is much more uncertain and questionable than that between the postulate of viewpoint-neutrality and the motive-based theory. This much seems to be common sense: while it is easy to attribute particular motives to a regulation which affects disproportionately one "viewpoint" more than another, it is much less clear whether a restriction of an entire subject-matter may be as easily traced to a specific type of motivation.

But is the assumption behind this whole argument plausible: is there really a strong link between a viewpoint-specific regulation and the likelihood of improper governmental motivations? To answer this question one must try to flesh out the notion of "improper governmental motivations". It is not good enough to dismiss viewpoint-based restrictions on the grounds that they reveal improper motivations unless we have a theory of what these improper motivations, in the context of freedom of expression, really are.

A standard answer is this: disfavour for a particular idea and, conversely, preference for another idea, constitute a class of improper legislative motivations: they must not activate the restriction of a particular utterance.[24]

---

21 See generally Stone, G R, "Content Regulation and the First Amendment", *Wm & M LRev* 25 (1983) 189; Stephan, P B, "The First Amendment and Content Discrimination", *Virginia LRev* 68 (1982) 203.

22 *Police Department v Mosley*, 408 US 92, 95 (1972).

23 Sunstein, note 20 above at 610; Stone note 21 above at 235; Williams, S H, "Content Discrimination and the First Amendment", *U Penn LRev* 139 (1991) 615, at 655-56.

24 Richards, D A J, *Toleration and the Constitution* (1986) 193; Alexander, L,"Low Value Speech", *Northwest ULRev* 83 (1989) 547, 553; Cass, R, "Commercial Speech,

But surely this is insufficient: law prohibits child pornography on the basis that the lawmaker disfavours the view which claims that exploiting children for sexual purposes is proper. To this claim, proponents of a "viewpoint-neutrality" theory may reply: the true motive of this particular governmental restriction is rather to prevent an identifiable harm to children. The problem with such a defence is that it is pedantic to distinguish between a "viewpoint" that a particular conduct is harmful and a judgment about the harm produced by that very conduct. A viewpoint-discourse is just another way of bringing the issue of harm into reasoning about restrictions of freedom of expression. If that is so, then we confront exactly the same problem as regards the previously discussed strategy: we might as well proceed directly to the speech-harm analysis, rather than disguise it in the language of a viewpoint analysis.

But there may be, naturally, genuinely improper motivations (even under the liberal conception of a limited government) which cannot be reduced to a judgment about harm produced by a given conduct, and which will taint a governmental restriction of free speech as illegitimate. An attempt to protect the executive branch against its potential critics, a preference for a particular religion or ideology, a desire to promote a particular type of moral virtue — these are examples of motivations which, if unrelated to the prevention of an identifiable harm, will taint a governmental action as illegitimate in the eyes of a liberal theorist. The point is that the postulate of viewpoint neutrality is too crude a device for distinguishing these motives from the motives that are related to an observable harm which the government has a duty to prevent (such as a "viewpoint" that sex with children is wrong, or that misleading advertising is improper). It may be easy to distinguish between viewpoint-neutral and viewpoint-based restrictions but the real work is done by a distinction within the class of viewpoint-based ones: between those restrictions which reflect improper motivations, and those which are activated by the proper ones. And if (as I have just suggested) this latter distinction cannot be made in isolation from a harm analysis, then an appeal to viewpoints is not helpful in a preliminary determination of the level of scrutiny to which anti-racial-vilification laws should be subjected.

This strategy, if successful, would offer more to the opponents than to the proponents of anti-group-vilification laws. This is natural: these laws seem to be clearly viewpoint-sensitive and so would trigger a rigorous test. But, for reasons just mentioned, I doubt the plausibility of the strategy, both on the basis of an observation that some perfectly unobjectionable laws do reflect a viewpoint-sensitivity, and also because it derives from the harm analysis. Ultimately, the difference between the second and the third strategies discussed in this part of the article seems to me insignificant in practice; it is no wonder that they share the same defect. While in theory one may distinguish between the analyses of the nature of speech (strategy (2)) and the motives for restriction of speech (strategy (3)), in practice the judgment about the latter seems to be inevitably based on the judgment about the former.

This is because, in practice, legislative motives are perceived as suspect when the nature of restricted speech is such that we expect the government to act out of concern for the communicative impact of the speech. In turn, this

---

Constitutionalism, Collective Choice", *U Cin LRev* 56 (1988) 1317, at 1352.

suspected governmental concern has the form of distaste for one viewpoint, or preference for another. We know from experience that in some areas this likelihood is less than in others, but in any event we make an inference about legislative motives from the character of speech in question. Those motives may be described as preferences for a viewpoint, or as assigning a lower value to a particular act of speech. In either case, the appeal to a broader theory (about the purposes which are served by freedom of expression, and which define what speech is central and what speech is marginal from the point of view of these underlying purposes) is inevitable. To this analysis we will now turn.

IV. The Relevance to the Principal Purposes of Freedom of Speech

In what became the most oft-quoted sentence in US First Amendment jurisprudence, Justice Stevens observed that "few of us would march our sons and daughters off to war to preserve the citizen's right to see 'Specified Sexual Activities' exhibited in the theaters of our choice".[25] Freedom of speech, though of great value, is not absolute, and some domains of freedom of speech are more valuable than others. To insist that all speech, regardless of its nature, should presumptively be equally protected, would be both patently absurd and inimical to the freedom of really "valuable" speech. Confronted with the demand for a uniform standard of protection of all speech, law would inevitably respond by lowering the degree of protection across the board.

As already observed, to insist that some types of speech are "central" while others lie at the periphery of the concerns of free-speech guarantees, immediately calls into question the nature of these concerns. The "centrality" of a particular type of speech lies in its connection with the underlying rationale for the special protection of speech. The problem lies, of course, in the inevitable indeterminacy of the underlying rationale of freedom of speech — disagreements about this rationale will affect disagreements about what types of speech deserve strong protection. The overview that follows is, of necessity, very sketchy. I shall not offer anything like a "general theory of freedom of expression". However, keeping in mind our subject-matter, I shall briefly review the main justifications for a strong protection of freedom of expression, and examine their implications for the argument about the degree of scrutiny that should be accorded to laws that prohibit group vilification.

A. Search for truth

The traditional Millian justification, taken in isolation from other rationales, supports a distressingly narrow scope for free expression. Most probably, it would defeat any call for a closer scrutiny of anti-racial-vilification laws — can you detect any contribution to truth in the mixture of silly stereotypes and straight lies in the opening paragraph of this article?

This, after all, may be a correct answer to our question about the standard of scrutiny, and yet we should pause before we accept this conclusion. The fact that the "search for truth" rationale would justify a high degree of censorship indicates that there is a clash between the dominance of this

---

25 *Young v American Mini Theatres, Inc*, 427 US 50, 70 (1976).

rationale and widely shared intuitions about freedom of speech. As Judge Easterbrook wrote: "[T]he Constitution does not make the dominance of truth a necessary condition of freedom of speech".[26] Nor is it the inevitable outcome of freedom of speech, though we might well hope that more often than not it will be the case.

To give some substance to this hope, we should distinguish among different types of speech, depending on their relevance to "the search for truth" (however indirect and remote this relationship may be). A crude distinction between statements of fact and expressions of opinion may be simplistic, and yet it indicates the type of argument worth considering here. Borderline cases notwithstanding, clear examples at both extremes of the fact-opinion spectrum suggest what principle underlies the distinction. "Cigarettes XXX contain 0.5 % tar" can hardly claim special protection, under the "search for truth" rationale, if demonstrably false. Indeed, we would expect the government to protect us against unfettered publications of lies of this type, and the same applies even more to, say, false attacks on private reputation. "There is nothing wrong with smoking, even if it shortens one's life: it is better to live happily, even if shortly" is the sort of opinion that is hardly amenable to the "truth" analysis. It comes close to the "pure opinion" end of the spectrum; we would not like the government even to start considering the conduciveness of this statement to truth, just as we would not want our law to decide for us the "truthfulness" of such statements as "Socialism is better than capitalism" (or vice versa).

The "pure opinion" argument is, at first blush, double-edged. It might justify a stronger protection for opinions — because they do not lend themselves easily to falsification, and so cannot be shown to be directly detrimental to the search for truth. But then they can be *denied* special protection precisely because they do not contribute to the search for truth. They simply do not lend themselves to truth analysis. Hence, depending on your perspective, Justice Powell's statement that "[u]nder the First Amendment there is no such thing as a false idea"[27] might as well be read as saying that "there is no such thing as a true idea". The operative words are, of course "under the First Amendment", and they indicate that the choice confronting us is political and legal, not epistemological. There are two possible ways out of this dilemma. One is to recite a presumption for non-restraint: it is restraint, not liberty, which calls for justification. In the present context, however, it becomes clear that the presumption itself must have *some* justification. Therefore there is another response: such a justification, and consequently a decision about the implications for the "search for truth" theory, must be found in rationales for freedom of speech other than the "search for truth" theory alone.

B. Self-expression

There is little doubt that human communicative activities are crucial for our capacity for self-expression and self-fulfilment. First, for self-expression, since speech is the most direct way of communicating to the rest of the society who we really are. The messages we convey to other people about our

---

26 *Hudnut*, note 7 above at 330.

27 *Gertz v Robert Welch, Inc*, 418 US 323, 339 (1974).

identity form part of a process that includes feedback from other people, which in turn modifies our own self-perception. Communication is in the centre of this complex process. Second, for self-fulfilment, since the exercise of our capacities is made possible only through self-definition, and the determination of who we really are is impossible without open communication with our fellow human beings. It also allows our self-fulfilment in a trivial sense: there are many of us who perceive our best capacities (rightly or wrongly) in areas having to do with communicating our ideas to other people.

But here lies the problem with the "self-expression" rationale for freedom of speech:[28] if it is derived from a general underlying principle of individual autonomy, then it must also share its limitations. If *the* reason for having freedom of expression is that it allows self-fulfilment, then self-fulfilment also determines the limits upon expression. And these limits are the same as in the case of any "self-expressing" or "self-fulfilling" conduct, viz. harm to others. Punching another person's nose may be the best way for you to express your real self, or to fulfil your true capacities, but we will not allow you to exercise your autonomy in this particular way. Extending it to our subject-matter: insulting racial minorities may be necessary for your sense of unrestrained self-expression and self-fulfilment. But it must be subjected to the same limits that we accept, in a liberal society, with regard to any conduct, communicative or otherwise. Hence an appeal to the values of self-expression and self-fulfilment is unhelpful to those who would apply a higher standard of scrutiny of laws that restrict freedom of speech. It does not follow that these values do not constitute an important part of the rationale for free speech. But it does follow that those who would insulate speech from the ordinary operation of the harm principle must find their rationale elsewhere.

C. *Tolerance*

The most original feature of an influential recent book by Lee Bollinger[29] is that it shifts the focus of arguments about freedom of speech away from the speakers and towards the audience. In Bollinger's view, it is not so much for the sake of protecting speakers but rather in order to promote right attitudes of tolerance among the audience that speech acts call for higher protection. "[T]he tolerance principle . . . is intended and designed to perform a self-reformation function for the general community and not . . . to offer a shield of protection either for the majority against the government or for minorities against unfair treatment at the hands of the majority".[30] Regardless of the lack of any social value in a given speech-act, promoting the attitude of tolerance towards the speaker helps to create "a general intellectual character" and enhances "the capacity for general tolerance".[31] Consequently, justified cases of non-protected speech, such as "fighting words", libel, or obscenity,

---

28 For examples of "self-realisation theory" of free speech, see inter alia Baker, C E, "Scope of the First Amendment Freedom of Speech", *UCLA LRev* 25 (1978) 964; Redish, M, *Freedom of Expression: A Critical Analysis* (1984) 9-86; Richards, D A J, "Free Speech and Obscenity Law: Toward a Moral Theory of the First Amendment", *U Penn LRev* 123 (1974) 45, 62.

29 Bollinger, L C, *The Tolerant Society: Freedom of Speech and Extremist Speech in America* (1986).

30 Id at 134.

31 Id at 182.

reflect those situations where self-restraint towards those speech acts cannot be reasonably expected.

This last issue is a good starting point for a critical reflection about Bollinger's theory. As it is, it fails to convince me that the underlying purpose of promoting tolerance can provide justification *both* for permitting "extremist speech" (which is the proper focus of his book) *and* for refusing protection to certain categories of speech that are outside the First Amendment protection. There is a distinct air of adhoc rationalisation in Bollinger's argument: taking the current Supreme Court's doctrine as a moral yardstick, he suggests that the protection of much of extremist speech can be justified by an appeal to the virtues of tolerance, while a refusal to grant higher protection indicates that it would be too much to expect of people that they tolerate some types of speech. One feels that this sort of argument might be used to justify any scope for protected speech, narrow or wide.

Perhaps this is somewhat unfair to Bollinger, since he attempts, at times, to explain why some instances of speech (for example, extremist political speech) lend themselves to improper intolerant responses more than other categories (for example, obscenity or libel). But at this stage, his explanations become anecdotal and of a very limited applicability.[32] More generally, Bollinger seems here to merge two ideas: a plausible (but not very original) one and a radical (but implausible) one.

The plausible idea is that even if much offensive or extremist speech is without any social value, restrictions are often demanded for wrong reasons, and those wrong reasons taint those regulations of speech. I have mentioned this idea before: while plausible, it does not offer any principled solution to our problem unless we have a theory about why certain types of restrictions are more likely to follow from wrong reasons, and therefore should trigger more careful scrutiny. To develop such a theory, we must focus on the value (or centrality) of certain types of speech that the government would like to suppress — hence we turn our focus back from the community's attitudes to the nature of the speech, and to the protection of the speaker. Bollinger's innovation is thus lost in the process.

The truly original idea suggests that promoting right attitudes among the audience, regardless of the protection of speech and/or of speaker, warrants a need for strong protection of some categories of speech, even if they are harmful under an ordinary calculus of harm. However, this sounds implausible. Under this theory, the law prevents A from interfering with B's speech in order to cultivate something important about A.[33] This sounds like a moralised paternalism: restricting an individual liberty for the agent's own good, where his or her good is defined in terms of promoting right virtues and commendablecharacter.[34] I do not think that such a reason for governmental action passes the muster of a liberal theory.

---

32 For instance, he asserts that much outcry against anti-Semitic marches and speeches in the United States is motivated by the attempt by Jews to appease their own sense of guilt, id at 129-130, 274-75 n17.

33 See, similarly, Schlag, P, "Freedom of Speech as Therapy" (Book Review), *UCLA L Rev* 34 (1986) 265, 279.

34 See Feinberg, J, *Harmless Wrongdoing: The Moral Limits of the Criminal Law*, vol 4 (1988) 299.

D. Democracy and self-government

Discussions concerning this rationale for freedom of speech invariably refer to the theory of Alexander Meiklejohn and his followers.[35] The argument is straightforward: democracy requires that citizens be free to receive all information that may affect their choices in the process of collective decision-making and, in particular, in the voting process. After all, the legitimacy of a democratic state is based on free decisions taken by its citizens regarding all collective action. Consequently, all speech that is related to this collective self-determination by free people must enjoy absolute (or near-absolute) protection. This applies not only to issues related to governmental processes, but more generally to issues of public concern, which in Meiklejohn's view extend to areas such as philosophy, literature and the arts.[36]

This last point has been taken up by a number of commentators who observed that a "democratic self-government" theory is either unduly restrictive (by providing protection only to political speech *sensu stricto*)[37] or virtually meaningless (because it broadens the meaning of the "political" so as to include in it every act of expression which deserves protection, regardless of its place in the political process).[38] Others have observed that self-government is not necessarily linked to the principle of strong protection of freedom of speech; indeed, one may perhaps argue *restricting* free speech on the basis of self-government.[39]

This observation would be damaging to the theory only if we adopted an unrestrained majoritarian view as the underlying conception of a democratic polity. On this view, democratically adopted limits upon free political speech would have to be respected and observed without dissent. And yet, the system of constitutional democracy, as we know and preach it, does not presuppose an unrestrained majority rule. Quite apart from all other restraints upon majority rule, which I shall not discuss here, at least this much seems obvious: the principle of democratic majority rule presupposes a free formation and expression of societal preferences and desires, and so it cannot condone any restrictions of channels of political communication. This is the theory which underlies famous footnote four of the *Carolene Products*[40] case: laws which tend to restrict the operation of democracy cannot be subject to an ordinary, democratic, majoritarian, deferential-to-legislator review, but must be subject to a stricter extra-legislative review.

---

35 Meiklejohn, A, *Free Speech and Its Relation to Self-Government* (1948); see also Bork, R, "Neutral Principles and Some First Amendment Problems", *Ind LJ* 47 (1971) 1; BeVier, L, "The First Amendment and Political Speech: An Inquiry Into the Substance and Limits of the Principle", *Stan L Rev* 30 (1978) 299.

36 See Meiklejohn, A, "The First Amendment Is an Absolute", *Supr Ct Rev* (1961) 245, 257.

37 See Tucker, D F B, *Law, Liberalism and Free Speech* (1985) 26-27.

38 See Tribe note 12 above par 12-1 at 786.

39 Bollinger note 29 above at 50-51.

40 *United States v Carolene Products*, 304 US 144, 152 n4. The relevant passage reads: "It is unnecessary to consider now whether legislation which restricts those political processes which can ordinarily be expected to bring about repeal of undesirable legislation, is to be subjected to more exacting judicial scrutiny . . . than are most other types of legislation". For a discussion, see Ely, J H, *Democracy and Distrust* (1980); Sadurski, W, "Judicial Protection of Minorities: The Lessons of Footnote Four", *Anglo-Amer L Rev* 17 (1988) 163; Ackerman, B, "Beyond *Carolene Products*", *Harv L Rev* 98 (1985) 713.

Note that in the absence of this (or a similar) theory, even if we adopted the democracy-based argument about freedom of speech, we would still not have solved the closer-scrutiny problem. It is, of course, true that free expression and communication are necessary for the healthy functioning of democracy, self-government and so on. But that is not the point of our inquiry: the point is to find out why this particular rationale would warrant the insulation of free speech from the ordinary operation of the harm calculus. In order to consider this issue, it is not enough to accept an uncontroversial thesis about the relevance of free speech to the democratic process; we must rather determine what it is about the democratic process that justifies subjecting free speech to a closer than usual scrutiny.

Consider again what "strict scrutiny" is about: it is really about a suspicion that politicians and legislators will overstate the degree of harm produced by a given activity (here: by exercises of free expression) and/or that they will undervalue the harm inflicted by the restriction. By subjecting a given restriction to a higher than usual standard of scrutiny we express our mistrust in the political judgment which affects the harm calculus. We must, therefore, have a theory about why such mistrust is more justified in one area than in the other. Why, for instance, is the law in a democratic community more likely to distort the harm calculus in the field of public concerns than in the area, say, of private defamation?

There are, roughly speaking, two plausible types of answers to this question, which will yield two types of theories of strict scrutiny of speech related to public issues. One theory refers to legislative motives, and the other refers to the risk of legislative miscalculation.

The motive-related theory stems from the observation that lawmakers tend to be self-interested and, hence, over-concerned about their own perceptions of harm. They will, understandably enough, view a strong criticism of their own action with dislike or suspicion, and they will perceive it as harmful to the public good. In reasoning about the legal restrictions that they propose, they will under-estimate the incidences of over-inclusion (that is, cases unnecessarily affected by a given restriction), but will be excessively sensitive to arguments of under-inclusion (which point at a need to expand a given restriction upon allegedly similar cases). In order to compensate for this self-interest motivated distortion, it is necessary to over-value the importance of free speech directed at public issues, and so to insulate it from the ordinary operation of the harm calculus.

The risk-oriented theory borrows from *Carolene Products*[41] the insight that legislative decisions, which may distort the future operation of democratic law-making, must be remedied through extra-legislative processes, and without the usual democracy-based deference to legislative judgment. Legislative restrictions upon freedom of speech, even if properly expressing current societal preferences, contain a high risk of distortion of preference-formation and preference-expression in future. In particular, they will make it more difficult for future lawmakers to remove these restrictions in accordance with changed public preferences because these restrictions will continually distort the overall picture of preferences. Restrictions upon free

41 See note 40 above.

speech tend not only to distort the process itself but also to distort the input of the process: lawmakers obtain a distorted picture about the actual distribution of various preferences and desires in a society. Hence, in a variety of ways, restrictions on public-concern speech pose a higher risk for the overall democratic process than many other restrictions do.

The upshot is that restrictions of speech trigger a higher standard of scrutiny of the speech-harm relationship if the speech at issue is related to public (and in particular, political) concerns. I do not wish to offer a definition of what makes a given subject-matter "removed from politics"[42] and what makes it closer to political and public concerns. I assume, however, that such distinctions *can* be made. Whether the criterion is that a particular type of speech "influences social relations and politics on a grand scale",[43] or concerns ideas which may "bring about... political and social changes",[44] is a matter for discussion. Our subject-matter (speech offensive to groups in a society) seems to fit clearly into this category.[45] This is a type of speech which affects the accommodation of interests among large groups in a community, which is meant to express a particular idea about the nature of the society in future (for instance, about its racial or ethnic composition), and which attempts to affect public policy (for example, the policy of immigration). That this speech is distasteful and scurrilous is another matter. But it *is* speech on public matters, and for this reason all restrictions on this speech must be subjected to a special, vigilant and suspicious, scrutiny, before the harm it produces warrants the prohibition.

V. Harms of Racial Vilification

There are three principal types of harm that result from offensive utterances made about a particular group in a society: (1) those that are produced by a violent reaction from the victims of group vilification; (2) those that result from violence committed by other persons, incited by "vilifiers" to assault the victims of racial vilification, and (3) those that consist in the very utterance of offensive words, regardless of any further violence committed. I will confine my further discussion to these three types of harm.

This is not to say that there do not exist other, collateral types of harm, which may be identified as ensuing from racial vilification. For one thing, one may argue that there is harm to the society as a whole, quite regardless of any more specific harm inflicted upon the direct targets of racial vilification. One might say that the broadcasting or publication of passages such as the one at the beginning of this article is harmful not only to those directly offended, but also to the community as a whole: it degrades the standards of everyday life, of civility, of "the quality of public discourse".[46] Further, it threatens social peace and harmony, and the society as a whole suffers as a result.

---

42 *Hudnut*, note 7 above at 331.

43 Id at 331.

44 *Roth v United States*, 354 US 476, 484 (1957).

45 See similarly *Anti-Defamation League of B'Nai B'Rith v FCC*, 403 F2d 169, 174 (DC Cir 1968) (defamation of a broad group or class "approaches the area of political and social commentary") (Skelly Wright, J, concurring).

46 Greenawalt, K, "Insults and Epithets: Are They Protected Speech?", *Rutg LRev* 42 (1990) 287, 302.

This may well be true, and yet I will not consider these harms as separate categories distinct from the ones I listed above. The reason is that these harms are either weaker than, or derivative from, harms that I consider to be the fundamental ones. A case for harm that consists in the lowering of the standards of civility in a given society is arguably weaker than the case for harm that consists in provoking violence, or inflicting psychic harm upon the target of vilification. It may well be that a society is prepared to abandon uniform standards of civility on the basis of, say, a libertarian view about unrestrained self-expression — but that does not dispose of the argument about harm inflicted upon the victims of vilification. At any rate, it is questionable whether a liberal society may enforce, through legal means, its views about the standards of civility.

As far as the breach of the peace and the threat to social harmony are concerned, this result derives from more specific harms listed above, namely, possible violence applied by the targets of vilification against the perpetrators or by third parties against the victims of vilification. These harms occur only if these more direct harmful effects occur in the first place.

A. Violent reaction by the victim

I witness someone giving a speech along the lines of the first paragraph of this article: the speech is deeply offensive to me, and I react violently. I try to shout the speaker down, someone in the audience is trying to carry me out of the room, and I push him back. A brawl follows. Who is to blame?

Morally speaking, the blame is probably shared (unequally) by the speaker, myself and those who took part in the violence. The speaker should have avoided the offensive language; I should have refrained from violent reactions; other members of the audience should have used peaceful means of solving the conflict between me and the speaker, and if these were insufficient, they should have called the police. But such an assessment of the whole situation is hardly relevant to the issue at hand. What we want to know is whether making an offensive speech should have been prohibited in the first place, by virtue of its tendency to cause a violent reaction by the victim of the speech. It is the speaker's right which is the focus of our attention at this stage, not the appropriateness of any particular responses to the violence following from the speech. And although in practical terms these two things may be hard to separate, for the purposes of a moral analysis of the speech-harm relationship we should concentrate on the former issue only.

While violence that is likely to follow from a group-vilification speech is undoubtedly a type of harm that a state is obliged to prevent, a strict-scrutiny analysis demands that the speech-harm relationship be direct and immediate, and that there be no other practical ways of avoiding the harm. Neither requirement is met in our hypothetical example. Violence follows directly from my reaction to the speech, not from the speech itself. If I simply left the room, or reacted in a peaceful way, the brawl might have been avoided. Consequently, there is an alternative way of preventing the harm, other than by silencing the speaker: it is to insist upon self-restraint and self-control by the targets of the speech. The question is, of course, whether this is a reasonable thing to do, for by such insistence we seem to be shifting the burden from the perpetrator to the victim.

To address this question one must introduce some further distinctions among different situations in which offensive words and utterances "assault" their targets. One distinction concerns the specificity of the assault: whether it is aimed at a particular person who happens to be in the audience, or is addressed to a group as a whole. This distinction will be considered in Section Three of this Part of the article. Another distinction concerns the avoidability of the assault. It is one thing to find oneself a member of a "captive audience", with no possibility of escaping a violent assault on one's reputation or dignity. It is another thing to suffer the sense of humiliation and ridicule where such an effect could have been expected, and avoided. This still does not redeem the offensive speech, but it significantly weakens the speech-harm link under the strict scrutiny analysis.

To express the idea that the hostility of the audience cannot justify suppression of the speech, the American free speech jurisprudence coined the concept of a "heckler's veto".[47] The idea is that to allow the hostility of the audience to warrant a restriction on the speech would amount to making the audiences the ultimate judges of constitutional rights. The end-result would be that we would "allow the intolerance (and threats) of a vocal minority (or even the majority) to determine who shall and shall not speak".[48] The law, as many liberals proudly remind us, is there to protect the speakers against intolerant audiences, not to protect the audiences against unpopular and even offensive speakers.

And yet, for all the persuasiveness of this principle, when applied to our subject-matter these words have a ring of unreality to them. Remember: we are talking about usually dispossessed, disempowered and poor minorities which are being ridiculed and offended by red-neck bigots and racists. Somehow, the picture of an insulted Pakistani youth in London or an Aboriginal in Sydney does not easily fit the concept of a "heckler" or a member of an "intolerant minority". On the other hand, a National Front or Klan leader does not immediately evoke the idea of an "unpopular" speaker who should be protected against the "hecklers".

Importantly enough, the doctrine of not allowing "hostile audiences" to cause suppression of unpopular speakers originated from a reverse fact-situation to the one exemplified by the typical cases of racial vilification. In the United States, the doctrine originally helped secure the speech of civil rights demonstrators in the South who protested against the conditions of discrimination in the early 1960's.[49] At that stage, the doctrine was fully compatible with the "fighting words" doctrine that refused protection to words that "by their very utterance inflict injury or tend to incite an immediate breach of the peace".[50] The "hostile audience" doctrine was associated with a value judgment about the content of speech in question, and consequently with the reasonableness (or otherwise) of audience hostility. It was only subsequently that the dissociation occurred, leading eventually to the use of the "hostile audience" argument in *Skokie* cases,[51] when the

---

47 The concept was initially used by Kalven, H, *The Negro and the First Amendment* (1965) 140-45.

48 *NAACP Legal Defense & Educ Fund v Devine*, 727 F2d 1247, 1261 (DC Cir 1984).

49 See Downs, D A, "Skokie Revisited: Hate Group Speech and the First Amendment", *Notre Dame L Rev* 60 (1985) 629, 632-36.

50 *Chaplinsky*, note 18 above at 572.

"hostility" of Jewish survivors of concentration camps to neo-Nazi ideas was not allowed to affect the right of neo-Nazis to march with Swastika signs throughout the Jewish populated suburb of Chicago. Now irrespective of all the arguments in favour of "viewpoint neutrality", we can surely draw a distinction between the intolerance of white bigots toward civil rights demonstrators, and the "intolerance" of Jewish citizens of Skokie towards the enthusiasts of gas chambers and concentration camps.

This doctrinal evolution (consisting in the expansion of the "hostile audience"-based protection upon worthless and distasteful speech) was logically combined with the gradual erosion of the *Chaplinsky* doctrine of "fighting words" ("logically" — because the *Chaplinsky* refusal to accord protection to "fighting words" could be maintained in its integrity only by confining the "hostile audience" doctrine to irrational, or hurtful, or illegitimate speech). In time, the thrust of the "fighting words" doctrine was mollified in a number of ways of which the most relevant for our purposes is the insistence that unprotected "fighting words" designate "captive" situations, in which the target has no reasonable means of escape.[52]

This seems to be crucial from the point of view of a strict scrutiny of the speech-harm relationship. For the question here is whether there are ways of avoiding or minimising the harm produced by speech related to public affairs other than by suppressing the speech itself. So far as we consider the dimension of harm related to the likelihood of violent audience reaction, the answer depends largely on the avoidability by the target of speech of finding oneself in a situation in which one's loss of self-control is excusable (say, under the standards of a defense of "provocation"). This applies to the situation in which offensive remarks assault a victim suddenly, and the target's reaction is immediate and spontaneous — but this does not extend to the situation in which a violent reaction could have been expected by the targets themselves, and avoided. That is why the Supreme Court of Illinois reminded the Jewish residents of Skokie that "it is their burden to avoid the offensive symbol [that is, the swastika] if they can do so without unreasonable inconvenience".[53]

B. Inciting others to commit violence against the victim

Adolf makes an anti-Polish speech which includes, and expands upon, the first paragraph of this article; Bruce, having heard the speech, repeatedly punches Czeslaw (you guessed it, Czeslaw is a Pole) in the nose and becomes a dedicated Pole-basher. Is Adolf responsible for this violence, and should this effect decide the legality or not of Adolf's speech?

Most people would probably say that it all depends, and that the criteria usually proposed for penalisation would concern the degree of directness (or "proximity") between the speech in question and the violent action, and also the existence of intent on the part of the speaker regarding the occurrence of

---

51 See *Village of Skokie v National Socialist Party*, 51 Ill App 3d 279, 366 NE 2d 347 (1977), *aff'd in part and rev'd in part*, 69 Ill 2d 605, 373 NE 2d 21 (1978); *Collin v Smith*, 447 F Supp 676 (ND Ill), *aff'd*, 578 F 2d 1197 (7th Cir), *cert denied*, 439 US 916 (1978).

52 See *Cohen v California*, 403 US 15, 21 (1971); see generally Tribe note above par 12-10 at 851.

53 *Village of Skokie*, note 51 above at 26. See, similarly, Cohen, C, "Free Speech and Political Extremism", *Law & Philos* 7 (1988) 263, 266-7.

these violent actions. People will disagree about whether both of these indicia must be present in order to make a speaker liable for the resultant harm, or whether either element alone constitutes a sufficient condition of liability. But two propositions will probably generate wide consensus. We cannot be held responsible for *any* harmful action anyone might take under the impact of our words — this would certainly stifle any free expression in the society. On the other hand, we cannot be absolved of *all* responsibility for the consequences of our words.

In order to chart a middle course between these two unacceptable extremes, one has to determine the requirements of "directness". One deceptively simple approach would be to apply a "but for" test: if it were not for the speech in question, violent actions would not have taken place. But the test is fatally difficult to apply with regard to the connection between a speech of one person and the action of another. There is simply too much that we (and the court) do not know about the subjective conditions of the agent; consequently, the only practical way to discern a causal connection between both is to inquire into the substance and the circumstances of the speech, and the mental propensities of the agent. But then the "but for" test virtually loses its meaning, for all we are saying is that a given speech was very likely to cause a particular response in a given person. In addition, the "but for" test would stifle a lot of speech which may be perfectly legitimate in most circumstances but, when heard by a person with unusually violent tendencies, may lead to aggression.

Another approach is to draw a distinction between different types of expression that motivate others to act. Thomas Scanlon draws an initially convincing distinction between speech that provides another person with the *means* to do what they wanted to do anyway, and expression that motivates others to act by pointing out what they take to be good *reasons* for action.[54] The example of speech of the first kind would be the disclosing to a prospective robber of the combination of the safe in the bank; the example of the latter, speech that provides arguments about why banks should be robbed. It is only in the former case, according to Scanlon, that the speaker may properly be held responsible for a harmful effect of another person's action. In the latter case, the responsibility of the agent who adopts the invidious reasons as his or her own supersedes, as it were, the agent's responsibility. "An autonomous person cannot accept *without independent consideration* the judgment of others as to what he should believe or what he should do",[55] and it is this "independent consideration" by the agent which is a direct source of a harmful action, thus "superseding" the speaker's arguments.

There is certain initial appeal in this proposal. It is based on a general liberal intuition about individual autonomy: everyone takes responsibility for his or her actions, and it is not the role of the government to protect mature, adult individuals from the influence of other people's pernicious ideas. But, while at first blush attractive, the principle based on Scanlon's distinction

---

54 Scanlon, T M, "A Theory of Freedom of Expression", *Philos & Publ Aff* 1 (1972) 204, 212 [hereinafter "Theory"]. Note that more recently, Scanlon retracted this part of his theory which hinges upon this distinction, Scanlon, T M, "Freedom of Expression and Categories of Expression", *U Pitt L Rev* 40 (1979) 519, 532.

55 Scanlon, "Theory", note 54 above at 216, emphasis added.

would lead to strongly counter-intuitive results. I keep suggesting to you, step-by-step and persistently, that there are excellent reasons for you to kill your neighbour who is a police officer. I impress upon you that society would be better without police, that the police are enemies of common people, that they are dangerous to all of us, and that the best way to get rid of this social evil is to kill each and every one of them. After a long and persistent indoctrination of this kind, you commit the murder of your neighbour. Am I not, both morally and legally, liable for the crime? I have provided you with *reasons* only, not with the means, and yet surely my responsibility cannot simply be "superseded" by your guilt consisting, as it does, of acting on such criminal and pernicious opinions.

One might say, of course, that a rational and basically moral person will not come to hold beliefs like those, even after continual persuasion. And yet we know that not everyone in a society is equally rational and moral, and we cannot found our principles about moral responsibility for evil done upon an abstract dogma about perfect rationality and basic morality of all adult individuals in society. Some people are more susceptible to aggressive ideas than others, and this fact must weigh upon our decisions concerning the exercise of our freedom of expression. Much of the judgment about this weight is context-bound. It is one thing to test arguments made during a university seminar in ethics about a duty to kill police officers (even though it would arguably be a very stupid thing to do), where we know and trust the good judgment of all our colleagues, and it is another thing to use these arguments in a speech addressed to a large crowd many members of which are unknown to us. In other words, we have a responsibility for anticipating the likely results of our acts of expression if these results may consist of violence done to other people, even if the connection between our expression and other people's violent action consists merely of providing the agents with the reasons to act in this particular way.

This example helps us to grasp not only the existence, but also the limits, of responsibility for inciting others to violence. There is a fine and yet crucial line between my telling you that police officers are evil, and my suggestion that you should murder a police officer. In the former case, I may be free of any murderous intentions towards police officers. I may, indeed, be strongly opposed to any violence against them. And the moral link between my anti-police views and your criminal action is quite remote: no necessary connection exists between those abstract views and your practical action. The difference here is the one between propagating a belief and advocating an action; in the words of an American judge, the "essential distinction is that those to whom the advocacy is addressed must be urged to *do* something ... rather than merely to *believe* in something".[56] But this is not enough because an expression of a belief may sometimes be a disguised form of the advocacy: the surface structure of speech in itself may be misleading. Something like a "clear and present danger" test[57] (or an equivalent) must therefore be introduced in order to judge the immediacy of the threat produced by one's words, even if these words sound more like a statement of belief.

---

56 *Yates v United States*, 354 US 298, 325 (1957) (Harlan, J, for the Court).

57 The test was initially formulated by Justice Holmes in *Schenck v United States*, 249 US 47, 52 (1919). For an application of the test in a modern landmark case, see *Brandenburg v Ohio*, 395 US 444 (1969).

These two criteria (advocacy of unlawful violence and immediacy of danger), applied independently of each other, seem to describe a plausible category of punishable incitement to crime. Note that this is not a strictly libertarian position: a libertarian would demand the joint application of the two principles, so that the absence of either would make the punishment for speech, or restriction of speech, unjustified.[58] And yet, in my view, it is plausible to say that advocacy of assault against other people should be penalised even if the harm is not imminent: this is derived from the prohibition of the principal crime of assault against others. On the other hand, even mere spreading of "beliefs" can be restricted if there is a clear and present danger to particular people and groups in society: this would satisfy a strict scrutiny test. But no spreading of offensive beliefs about groups should be prohibited (under this category of harm) if no imminent danger of violence is ascertained.

Ultimately, the argument for penalising group vilification on the basis of its tendency to induce others to act violently cannot be based on incitement, ie, on such advocacy of harm where the connection between speech and harm is proximate and straightforward. These situations are handled by imposing penalties for inciting others to commit a crime;[59] there is no need for a new criminal category there. Penalising group vilification applies therefore to situations that are not captured by "incitement", thus to those cases where the connection between speech and another person's criminal behaviour is more remote.

But this is a dangerous consequence because where the danger is not "clear" and/or "present", it is both unfair to a speaker and stifling to the general exchange of views in the society, to hold the speaker responsible. Such responsibility might perhaps be detected under a very lenient standard of scrutiny. We might say that widespread racist propaganda may, in the long run, contribute to an increased level of racist violence in the society. But under strict scrutiny, which calls for the determination of a clear, immediate connection between the prohibition of conduct and prevention of an identifiable harm, group vilification which falls short of incitement to crime cannot be prohibited on this particular ground.

C. *Psychic injury*

As a recent migrant from Poland, who wants to be an active and loyal citizen of his new country, I go to a meeting of the town council in my suburb. At the first meeting, quite unexpectedly, a speaker rises and makes a fiery speech in which he proposes a zoning ordinance that will make it more difficult for new immigrants, especially for Poles, to settle in the suburb. This, he says, is to protect the safety of children and the standards of cleanliness in the suburb. The argument runs along the lines of the first paragraph of this article.

I do not react violently to the speech (incidentally, the motion was lost by a wide margin): I am too frightened, too intimidated, or too embarrassed about

---

58 See discussion in Redish, note 28 above at 81-6, 197-9.

59 In Australian and English law, for instance, a person who counsels, commands or advises the commission of an indictable crime is guilty of the common law misdemeanour of "incitement", see *R v Quail* (1886) 4 F & F 1076, 176 ER 914 (inciting to larceny); *R v Howe & Ors* [1987] 2 WLR 568 (inciting to murder); *R v Ranford* (1874) 13 Cox CC 9 (attempt to incite).

my foreign accent, or perhaps it is simply that violence is not my way of reacting to insults. No-one else assaults me in any perceptible, physical way. And yet, since the day I have heard the speech, my life has clearly been transformed for the worse. Whenever I meet my neighbours, fellow workers, or salespersons in the shops, I search for expressions of dislike or contempt in their eyes. When they are rude, I attribute it to their hatred of Poles. When they are polite, I treat it as a symptom of their patronising attitude, or their protecting me against distress. They know that I am Polish. I know that they know. And they know that I know that they know.[60] From now on, my relationship with others can no longer be normal, because we are entangled in a web of mutually self-reinforcing attitudes of suspicions, fears and dislikes. My self-image is inevitably shaped by my belief that others are contemptuous of me. I feel schizophrenia developing about my attitudes towards my own identity: at the same time, I am proud of my heritage and wish to get rid of the difference that costs me so much in this country. I am upset, distressed, angry and deeply offended.[61]

Liberal political philosophy is full of arguments about the insufficiency of mere offence in warranting the prohibition of an offensive expression. Liberal jurisprudence proudly declares that offensiveness of a speech-act is not reason enough for restricting it: "the public expression of ideas may not be prohibited merely because the ideas are themselves offensive to some of their hearers".[62] Indeed, some liberals go a step further and claim that the fact that a given speech is offensive or distressing is all the more reason to protect it against restrictions.[63] And yet, for all the attractiveness of the general principle, and for all the talk about protecting unpopular views against majority tyranny, there is a strange lack of correspondence between these high liberal proclamations and such personal experience of hurt and humiliation as described above. For I would genuinely be surprised if told that the damage done to me by racist speech which humiliates my group can be summarily described as "mere offence".

Liberals traditionally have had trouble grasping the severity of, and relating sympathetically to, the kind of psychic harm inflicted by group vilification. A possible reason for this is that they have usually applied to psychic harms the imagery of members of the majority offended by unorthodox, shocking minority expressions. Within this imagery, the power of the state to protect people against harms to their sensibilities would indeed "effectively empower a majority to silence dissidents simply as a matter of

---

60 In a classic discussion of the relationship between "the stigmatised" and the "normal people", Goffman described this mechanism as "the infinite regress of mutual consideration", Goffman, E, *Stigma* (1968, 2nd edn) at 30.

61 On racial stigmatisation, see Lawrence, C R, "The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism", *Stan LRev* 39 (1987) 317, 349-55. For an important study of the effects of ethnic/racial prejudice and vilification upon its victims, see Simpson, G, & Yinger, J, *Racial and Cultural Minorities: An Analysis of Prejudice and Discrimination* (4th edn 1972), esp at 142-64. See also Davis, P C, "Law As Microaggression", *Yale LJ* 98 (1989) 1559, 1565-68.

62 *Street v New York*, 394 US 576, 592 (1969).

63 "[T]he fact that society may find speech offensive is not a sufficient reason for suppressing it. Indeed, if it is the speaker's opinion that gives offense, that consequence is a reason for according it constitutional protection", *FCC v Pacifica Foundation*, 438 US 726, 745 (1978).

personal predilections".[64] Consequently, some liberals point out that "if 'offensiveness' were the test, majority rule would replace the first amendment".[65] But the pattern of the majority versus minority clash in the racial vilification laws is quite different: it is not a case of the majority trying to silence the minority but rather of a minority trying to silence the intolerant majority, or trying to enlist the support of the majority in silencing a vicious minority. The silencing involved in enforcing anti-racial-vilification law is not the kind of silencing associated with majoritarian oppression. A minority that seeks help through anti-racial-vilification laws is precisely the sort of group which has been traditionally seen by liberals as deserving special legal protection against possible majoritarian oppression: a powerless, subordinated and disadvantaged minority. So the whole pattern of minority protection versus majoritarian oppression is, in our subject-matter here, opposite to the one which prompts many liberals to raise the alarm against granting legal recognition to one's aversion towards others' speech.

In contrast to the orthodox liberal view, I will take as obvious that offensive remarks about one's group do inflict harm upon group members. Sometimes. The questions are: whether, and why, this harm is different from the sort of harm that leads unobjectionably to the legal prohibition of any other harmful conduct. It is important to note, first, that the issue at hand is not just *any* offensiveness of a speech-act. Whether words and images that harm individual sensibilities because they are shocking should be legally prohibited is not the question here. We are concerned with a *particular* kind of harm to one's sensibility; that is, the harm resulting from vilification of one's own group. This is a different matter from someone claiming protection from unwanted exposure to disgusting words and images. The difference is in the degree of implication of one's own identity. If someone protests against unwanted exposure to pornography or indecent language, then one is more remotely and indirectly implicated in the subject-matter of the offensive material than when one protests against contemptuous or hateful statements concerning one's own racial or ethnic group. It is a matter, so to speak, of having a moral standing to protest. This is important to stress because some liberals have a tendency to subsume group-vilification actions under a broader category of offensiveness, and then easily dispense with the problem by appealing to the principle that offensiveness alone is insufficient ground for restriction. A non sequitur arises in the argument because group vilification is not a matter of "mere offensiveness".

But the matter is, unfortunately, more complicated than that. Forget racial vilification for a moment and think of some other types of group vilification. Blasphemy is an example of speech that is either "a mere offence" or a "group vilification", depending on the hearer's perspective. If I identify with my religion strongly (which is, after all, what all religious adherents are supposed to do), then it means that offensive remarks about my religion do not merely "offend" me; they implicate my own identity in a way which expresses contempt, humiliation, and hate towards me. To say that a Muslim is "offended" by *Satanic Verses* does not fully capture the effect of Rushdie's

---

64 *Cohen v California*, 403 US 15, 21 (1971).

65 Shiffrin, S, "Defamatory Non-Media Speech and First Amendment Methodology", *UCLA L Rev* 25 (1978) 915, 951.

novel upon the psyche of Muslims: it hurts them (I imagine) personally because they see themselves implicated in a story which (they think) expresses contempt for their group. And since their group moulds their own identity to a high degree, they feel hurt themselves. So in the case of pornography: to some viewers it may be merely offensive, but to those women who see it as expression of contempt for women, or part of an ideology that treats women as objects of sexual exploitation, it is more than a "mere offence": it is an insult to themselves in a way which implicates their own identities.

My point is not simply that we have here a line-drawing problem between "mere offence" and offence that implicates the identity of a viewer/hearer. The problem is more serious: whether we classify a given offence as falling on one or the other side of the distinction depends ultimately upon our substantive assessment of the merits of the vilifying speech. And this has some disturbing consequences.

Consider first the issue of "symbolic speech". According to some commentators, acts such as wearing a jacket with words "Fuck the Draft"[66] cannot be prohibited because, in contrast to group vilification, such acts do not harm anyone. In the words of a writer who is otherwise in favour of criminal liability for group vilification, symbolic offensive speech would escape punishment because "an essential characteristic of symbolic speech . . . is that *by definition* it poses no serious harms to substantial public interests".[67] This is supposed to contrast with harm caused by "fighting words" and other categories of non-protected speech. But the distinction is ultimately in the eyes of the beholder: one may well imagine a situation in which harm to one's psychological well-being and self-respect, caused by a "Fuck the Draft" sign, may be serious and (more importantly) may implicate one's own identity (think of a disabled war veteran, or parents of a soldier who is currently participating in a war, etc). Hence, the judgment that no serious harm to one's psyche is caused by a particular offensive sign hinges upon a judgment about the substantive value of this sign, and more importantly, about whether the viewer *is justified* in finding his own identity implicated by the sign.

Another example: in a recent Note which advocated penalising group vilification, the *Harvard Law Review's* editor(s) suggested that "the law should recognise the sensibility harm experienced by the targets of group vilification",[68] which is distinct from "interest harms" (the sole form of harm recognised by the courts so far). The difference is that in the case of "sensibility harms" the harm cannot be defined independently of the hearer's attitude towards the speaker's point of view, while in the case of "interest harms" it can. The Note's recommendation for such an extension appeals directly to a communitarian vision of individual and society: "the communitarian argument holds that the law should recognise the sensibility harm experienced by the targets of group vilification, because to disregard it

---

66 See *Cohen v California*, 403 US 15 (1971).

67 Note, "Group Vilification Reconsidered", *Yale LJ* 89 (1979) 308, 322 n59, emphasis in original.

68 Note, "A Communitarian Defense of Group Libel Laws", *Harv L Rev* 101 (1988) 682, 692.

is to compromise the shared commitments that make freedom possible".[69] So far so good. The problem is that if we allow all "sensibility harms" to be protected by law, we shall end up restricting even the most valuable and justified speech, as long as the sensibility of a criticised person is harmed. What about civil rights demonstrations which may offend the sensibility of a racist? To escape such counter-intuitive conclusions the authors of the Note draw a further distinction between "instrumental" and "constitutive" communities; it is only in the latter that members' identities are directly affected by the statements related to the community as a whole.[70] Most communities, however, are instrumental, and so do not warrant sensibility-harm based protection against group vilification. For instance, a white bigot will not be able to claim protection of his "sensibility harm" against civil-rightists, because the "communities" affected by anti-discrimination speech, such as the Ku Klux Klan, are not constitutive.

The aim of the argument is clear: it is to reconcile a prohibition of group vilification with the denial of such a protection to wrong groups. But the argument only pretends that it draws a neutral and clear line between "constitutive" and "instrumental" communities; in fact, the operative line is between those communities of which we approve and those of which we do not. Consequently, the distinction has nothing to do with the seriousness of the implication of one's identity in the group of which one is a member. A Ku Klux Klan member may well be psychologically and morally affiliated with his organisation to a higher degree than are many other people affiliated with their nations or their religions. His identity may be shaped by his Klan membership to a very high degree indeed. And yet we rightly deny him protection from criticism of the Klan. We do it not because we doubt the seriousness of his identification with the Klan but because we believe that the Klan's ideas are not worthy of protection. Hence, the distinction between "constitutive" and "instrumental" groups is just a proxy for a substantive judgment about those group identities that deserve protection and those that do not.

But if this is the case, then it prompts us to rethink our initial point about a distinction between "mere offensiveness" and offensiveness that directly implicates one's own identity. This distinction disappears unless we are prepared to say openly that it hinges upon our views about which group identities are worth protecting and which are not. If we *are* prepared to engage in such a value judgment about the worthiness of some sensibilities that deserve protection, then we must face the consequences of unrestrained majoritarianism. For example, we will have to say that we disallow Muslim claims for special protection against blasphemy because we value their religion less than other religions. And the same will have to be said to women who want protection from "sensibility harm" caused by pornography: their harm, even if genuine, will be seen to be less worthy of protection. I doubt if any legislator or legal theorist would be happy to accept this consequence.

The only alternative to drawing the distinction (between a mere offence and an identity-implicating offence) in this way is to abandon the distinction altogether, and to lower the level of legal protection for sensibility harm

---

69 Id at 692.
70 Id at 693-94.

across the board. This, I believe, is a more candid and honest solution. The theory behind it may be summarised as follows: The severity of sensibility harms is in the eyes of beholders. Law does not draw the line between sensibility harms worthy of protection and those less deserving of legal concern. The degree of one's personal identification with his or her own wider group is not determined by the law. There is, therefore, no such thing as an agent-neutral, rigid distinction between "mere offence" and offence that implicates a person directly through insulting his or her group. *But* the price to be paid for this legal abstention from judgments about offensiveness is that, in order to win legal protection, a claim for prohibition of group vilification must pass strict scrutiny of the speech-harm relationship. This is the trade-off: all complaints about vilification are considered genuine (that is, as implicating one's own identity through vilification of the group), but since a discourse about groups is close to a public discussion about political matters, in a self-governing democratic society it must be insulated against an ordinary calculus of harm.

This indicates the main difference, from the point of view of our analysis, between individual defamation laws and laws which punish for group vilification. Even if both types of speech, covered by these laws, inflict similar injury upon the psychological well-being of the victim, individual defamation is so remote from discourse about public and political affairs that it may be measured by ordinary scrutiny without any risk to self-government: the harm inflicted must simply be compared with the harm of restriction. But in the case of group vilification, the threshold of harm (to be demonstrated as produced by an act of speech) must be much higher because the speech in question is more relevant to the debate about public matters in a community.

This is not to say that racial vilification laws will necessarily fail the test, but that the test of harm must be much more stringent than the one applied, say, to defamation laws. The stringency lies, not merely in requiring a higher-than-usual demonstrated seriousness of harm, but also in requiring proof that this kind of prohibition is the only way to avoid the harm.[71] A state education policy and the enforcement of anti-discrimination laws spring immediately to mind in this context. If we consider the availability of such alternative ways of reducing emotional harm, and also that harm is unlikely to follow from every speech-act that (under the racial vilification laws) would qualify for restriction, the harm of psychic and emotional injury is unlikely to overcome the hurdles of a strict scrutiny.

I wish to emphasise that mine is a different position from the standard liberal argument that attempts to show that group vilification is somehow less severe, less serious, more "diluted", than an individual insult or defamation.[72] It is said also that in the case of group defamation, the "nexus" between defamation and the complainant is less certain than in the case of individual libel.[73] In traditional legal argument it is established that "to be actionable the defamatory words must be understood to be published *of and concerning* the

---

71 See, similarly, Heins, M, "Banning Words: A Comment on 'Words That Wound'", *Harv CR-CL L Rev* 18 (1983) 585, 586-7.

72 For a summary and criticism of this argument, see Note, note 67 above at 310-12.

73 See for example, Note, "Group Defamation: Five Guiding Principles", *Texas L Rev* 64 (1985) 591, 596 n26.

plaintiff"[74] and this condition is easier to meet in individual rather than group defamation cases. And yet, regardless of the evidentiary problem of demonstrating the "nexus", I do not think that one can argue that group vilification is in general less severe than individual insult. It may well be that some racial slurs against my race will concern me less than insults attacking me personally, but one can very well conceive of the reverse situation: I may be more hurt by an insult against me *qua* member of my group than *qua* individual. For this reason, Justice Frankfurter had a point in attacking in *Beauharnais* the view according to which "speech concededly punishable when immediately directed at individuals cannot be outlawed if directed at groups with whose position and esteem in society the affiliated individual may be inextricably involved".[75]

These words have an obvious ring of truth, especially if one considers the specific subject-matter of the *Beauharnais* case, viz. punishment of Joseph Beauharnais for distribution of particularly odious racist leaflets addressed against Black citizens of Chicago. From the point of view of the severity of injury, Mr Beauharnais' leaflet must have been at least as hurtful to a Black reader as insulting words shouted directly to him or her. And yet the difference lies in the rationale for a standard of scrutiny: Beauharnais' appeal to the authorities of Chicago to segregate Blacks, odious though it is, is his contribution to public debate about racial relations in his country. Insulting words addressed to an individual Black are not.

This, I believe, is a good reason for justified suspicion towards group libel or group defamation laws. In American law, for example, it has long been established that criminal libel has a scope limited to defamation of individuals, not groups.[76] Similarly, unprotected "fighting words" have been clearly confined to face-to-face situations.[77] These are the circumstances that undermine virtually any claim of these utterances to a political role, to a function in the debate about public issues. But the situation is different with regard to statements, however odious, about large groups.

VI. Conclusions: Liberalism and Group Vilification

"Even though members of the self-appointed liberal elite would never dream of stooping to racist speech, neither, it seems, would they ever dream of taking legal steps to stop it".[78] These harsh words by an English lawyer and writer, no radical himself, indicate how serious a dilemma is raised for liberals by proposed and actual laws that restrict racist speech. This is perhaps

---

74 *Knupffer v London Express Newspaper Ltd* [1944] AC 116, 121 (HL) (emphasis added).

75 *Beauharnais v Illinois*, 343 US 250, 263 (1952).

76 *Beauharnais* was the Supreme Court's only decision that upheld (by a bare majority) the validity of a group defamation law. Since then, a number of decisions implicitly undermined the authority of *Beauharnais* (see, inter alia, *New York Times v Sullivan*, 376 US 254 (1964), and "Skokie" decisions, see note 51 above), and it is generally considered that *Beauharnais* is no longer good law.

77 See *Chaplinsky* note 18 above at 573; *Texas v Johnson*, 109 SCt 2533, 2542 (1989). On racial insults directed at particular victims, see Delgado, R, "Words that Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling", *Harv CR-CL L Rev* 17 (1982) 133, 174-5.

78 Lee, S, *The Cost of Free Speech* (1990) 42.

the most serious challenge to liberal orthodoxy which has prevailed in the West with respect to the philosophy of freedom of expression.

This liberal orthodoxy has been based on a radical severance of the moral assessment of conduct from the determination of one's freedom to engage in it. The former is guided by the argument of "the good", the latter by the argument of "the right". The cherished regime of rights precludes the operation of moral judgments of the good in the area of the right; in this area, the state must be neutral between competing conceptions of the good. Hence, in an apt phrase of Michael Sandel's, "[l]iberals often take pride in defending what they oppose — pornography, for example, or unpopular views".[79]

But there are limits to one's capacity of "taking pride in defending what one opposes" — and beyond these limits, schizophrenia looms large. You can persist in defending individuals' rights to reprehensible conduct as long as the conduct in question does not bother you too much, but as the severity of your condemnation grows, the intellectual and psychological resources necessary to protect a right to this conduct get exhausted. In other words, the conflict between "the good" and "the right" is reasonably manageable as long as the violation of "the good" is not of dramatic proportions.

That is why a liberal has more problems in defending a right to racist speech than, say, a right to flag burning: the evil of the latter seems to us more symbolic than substantive in terms of hurting identifiable individuals. But even with regard to these rather symbolic provocations and offences, patience towards the liberal argument of the sort, "I hate what you are doing but I respect your right to do it" seems to be not unlimited. In her recent remarkable article, Robin West expresses precisely such an impatience with the attitude of many American liberal lawyers to the recent controversy about flag burning cases and the proposed constitutional amendment.[80] In testimony before the Congress and in various publications, prominent liberals "made clear that their defense of the rights of flag burners did not rest on agreement with the ideas or the individual expressing the ideas. If anything, liberals went to excessive lengths to *dissociate* themselves from the content of the flag burners' protest".[81] This pattern of "defending what one opposes" is consistent with traditional liberal orthodoxy and yet, according to Robin West, it is ultimately formalistic and possibly counterproductive. It defends rights for rights' sake (rather than for the sake of values which they promote), and it denigrates the social dissent expressed in such actions as flag burning. The alternative path to defend the right of flag burners, West argues, is by appealing to their functions in a socially valuable process of communication and self criticism. "The measure of our liberality becomes not the extent to which we can tolerate the offensive, hateful, or simply unpopular ideas of others, but the degree to which we individually take responsibility for the truth of our utterances and collectively value and nurture the communicative and truth-promoting realms of political culture and political dissent".[82]

---

79 Sandel, M, "Introduction" in Sandel, M, (ed), *Liberalism and Its Critics* (1984) 1.

80 West, R, "The Supreme Court, 1989 Term — Foreword: Taking Freedom Seriously", *Harv L Rev* 104 (1990) 43

81 Id at 94-95, footnote omitted, emphasis in original.

82 Id at 96, footnote omitted.

This path of discerning positive values in actions that initially seem merely offensive and provocative, may be available to us with regard to flag burning but not with regard to group vilification. And yet West's general observation is instructive: it may be simply not good enough to defend a right to a practice in which we can discern absolutely no redeeming positive features. And so, at the end of this essay, it may perhaps be necessary to reflect upon whether the practice of allowing group vilification to remain unpunished has some benefits of its own (even if group vilification itself is uncontroversially odious and harmful, as it no doubt is).

This has *not* been the general strategy of this article. Rather, the strategy was to show that group-vilifying speech warrants a high level of protection in terms of the speech-harm test, due to the proximity of that speech to a public debate on political issues. Further, it was suggested that laws prohibiting racial vilification do not pass this test in any of the three main areas of harm: harm resulting from violence by a target of the speech; harm resulting from violence by other people against the target; and harm inherent in the offensiveness of the speech itself. But the assumption underlying the argument was that group-vilifying speech *is* harmful, and that the harm it inflicts is such that it is proper for the state to treat it as a morally relevant reason in support of proposed legal restrictions of liberty. It is on the judgment about how this harm fares, in the strict scrutiny of laws that restrict speech, that the burden of my conclusion against anti-group vilification laws rests.

But now that the principal argument has been submitted, we may go a step forward, lift an underlying assumption, and, in accordance with Robin West's remarks, consider whether there may be any good in allowing racial vilification to go unpunished. That is to say, is there any social good, other than the underlying value of strong protection of free speech, related to self-government in a democratic state?

First, one good of legality of such speech is that it might prevent a complacent attitude of society towards the existence of racism and racist attitudes. Racists *are* there, and it is better to let them air their views in the open rather than allow an illusion to grow that the problem has been solved because racist statements have been made illegal. Group vilification is a symptom, not a source, of deeper problems that give birth to hate and contempt by some people for other groups in society. By prohibiting public statements that vilify those groups we may slightly reduce the hurt to the feelings of their members, but at the same time we risk removing the issue of racism from the public agenda. The good of allowing group vilification is that it helps maintain the visibility of a dramatic problem which is there anyway, regardless of the prohibition.

Second, legality of group vilification may, ironically perhaps, be valuable for the subordinated groups themselves. In a recent article, Kenneth Karst points out that unrestricted freedom of expression is a mixed blessing for the members of subordinated groups: on the one hand, they are victims of speech by members of the powerful groups in their society, but on the other hand "precisely because an important part of a group's subordination consists in silencing, their emancipation requires a generously defined freedom of expression, a freedom that overflows the shallow capacity of the model of civic deliberation".[83] I am not sure which benefit prevails in the calculus of

costs and benefits for the subordinated and powerless groups, nor am I sure whether such a calculus can be made in abstract terms. And yet it seems to me that a broad regime of freedom of expression, which allows speech considered insulting, offensive and shocking to remain unpunished, may be at times useful to those groups who are alienated from the cultural and political mainstream.[83] Conversely, it is likely that racial vilification laws *may* be invoked (even against the best intentions of their drafters) to silence the anti-establishment speech by disadvantaged groups, who might have to resort to shocking and "uncivil" expressions: to overcome the systemic bias of official channels of communications, and to get their message across to the community at large (after all, "civility" is defined by a dominant culture which has all the conventional means of mass communication at its disposal). Remember the warning by Justice Black when he dissented from the *Beauharnais* decision which upheld an anti-hate-speech statute: "If there be minority groups who hail this holding as their victory, they might consider the possible relevancy of this ancient remark: 'Another such victory and I am undone'".[84]

Third, legal tolerance of group vilification has expressive value contained in the message it conveys to people about the nature of their society. To describe it in deliberately extreme and exaggerated terms, the message is as follows. In a liberal society claims to protection against insult and offense are viewed with utmost suspicion. Racial vilification approaches the outer limits of the rationale for tolerance, but this is precisely why the educational effect of this tolerance may be so powerful. Liberal society is opposed to a communitarian vision of legally protecting an individual's sense of identity with a wider group; under that vision, an attack on the group is viewed as an attack on an individual. In contrast, in a liberal society people are encouraged to distance themselves from their collective identities, to treat them as social roles rather than as ingredients of their own selves, and to put up with many dignitary injuries that other societies would have treated much more seriously. To use a traditional distinction, a liberal society is much more of a *Gesellschaft* than a *Gemeinschaft*, or in a felicitous term recently coined by Meir Dan-Cohen, "a union of detached roles".[85] As contrasted to the ideal of "community", in such a union people are encouraged never fully to identify themselves with the roles they occupy, and their identity is never fully defined by the membership in groups. People are urged to be able to stand back from the roles they are playing, to adopt a critical attitude to them, and to perceive their own identity as transcending the sum of roles stemming from their involvement in various groups. This is one practical sense of the controversy about "thick" (ie, context-bound) versus "thin" (or "unencumbered") conceptions of the self in a liberal society.[86] It would not

---

83 See, similarly, Partlett, D, "From Red Lion Square to Skokie to the Fatal Shore: Racial Defamation and Freedom of Speech", *Vanderb J Transnat L* 22 (1989) 431, 477. See also, more generally, Crenshaw, K W, "Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law", *Harv LRev* 101 (1988) 1331, 1352-69, 1381-84.

84 *Beauharnais* note 7 above at 275 (Black J, dissenting). See also, in the same decision, a similar warning by another dissenting judge: "Today a white man stands convicted for protesting in unseemly language against our decisions invalidating restrictive covenants. Tomorrow a Negro will be haled before a court for denouncing lynch law in heated terms." 286 (Douglas J, dissenting).

85 Dan-Cohen, M, "Law, Community, and Communication", *Duke LJ* (1989) 1654, 1669.

be easy to have "enough self-control to refrain from violent responses to odious words and doctrines"[87] if these "words and doctrines" were seen as destroying one's central aspects of self. But it can more easily be done if these insults are not allowed to reach your own identity: when they attack only some of your social roles, which are seen as extrinsic to your real self, you may easily "turn on [your] heels and leave the provocation behind".[88]

A determination of the degree to which your identity is constituted by your community involvements is, from this perspective, just another *facon de parler* about how to react to unpleasant or even odious utterances by others. A pluralistic, heterogeneous liberal society, must lift a number of traditional protections of one's psychic well-being, so as to maintain its pluralistic and cosmopolitan character. In a sense, a liberal society rejects the principle of honour as a good which one may protect through law; the idea of honour related to the community-defined individual is replaced by the central conception of an autonomous individual who may shape his or her social identities and attachments.[89] Individual dignity is perceived more in the power of autonomous shaping of one's own social world rather than in the existence of a protected sphere of communal attachments which mould an individual self. Hence, this liberal insensitivity to many psychic harms is the price of a broadened scope for individual autonomy.

Many will think that the price is too high. Many will propose to draw the line between protection of individual dignity and protection of individual autonomy at another point, so that group vilification will fall on the prohibited side of legal restrictions. My main argument against penalising group vilification does not hinge upon this expressive value of tolerating such speech. But this last point suggests that, whatever our substantive judgment about legality of these restrictions is, the importance of the problem is that it prompts us to reflect upon the relations between the scope of legal prohibitions and the vision of the society we want to live in.

---

86 Compare Sandel, M J, *Liberalism and the Limits of Justice* (1982) 150; Walzer, M, *Spheres of Justice* (1983) 8, Goldman, A H, "Real People (Natural Differences and the Scope of Justice)", *Canad J Philos* 17 (1987) 377, 382-3 with Rawls, J, "Justice as Fairness: Political not Metaphysical", *Philos & Publ Aff* 14 (1985) 223, 234; Gray, J, "Contractarian Method, Private Property, and the Market Economy", *Markets and Justice: Nomos XXXI*, Chapman, J W, & Pennock, J R, eds (1989) 39; Sadurski, W, "Natural and Social Lottery, and Concepts of the Self", *Law & Philos* 9 (1990) 157, 172.

87 Feinberg, J, *The Moral Limits of the Criminal Law*, vol 2: *Offense to Others* (1985) 91.

88 Ibid.

89 For a contrast of autonomy and reputation-as-dignity, see Post, R C, "The Social Foundations of Defamation Law: Reputation and the Constitution", *Calif LRev* 74 (1986) 691, 731-9.